Title
Habit Formation In Exercise: An Empirical Analysis of Exercise Habits

Author
Dominguez, Jorge

Publication Date
2016-10-24

Habit Formation In Exercise:
An Empirical Analysis of Exercise Habits

Jorge Dominguez
Mentor: Dr. Heather Royer
Department of Economics
University of California, Santa Barbara

Abstract

This paper analyzes habit formation in exercise by examining the interdependence between past, present, and current exercise levels. Using login record data from a gym located in the Midwest, along with Quality Controlled Local Climatological Data from the National Oceanic and Atmospheric Administration (NOAA), we estimate the effects of snowstorms on gym visits. The empirical results indicate that severe snowstorms have a causal effect on gym attendance. In light of these results, we use snowstorm variation as an instrument to estimate the dynamic relationship between past, present, and future exercise. We find that past exercise behaviors have an influence on present habits, implying that exercise routine interruptions may have a strong negative impact on habit formation, and that current exercise behaviors depend on anticipated future exercise. Our results imply rational habitual behavior in exercise and provide new insights for the study of habit formation.

1. Introduction and Literature Review

A. *Introduction*

Sedentary lifestyles have become increasingly common among U.S. adults (Blumenkranz, Garber, and Goldhaber-Fiebert, 2010). According to the Centers for Disease Control and Prevention (CDC), nearly 80% of adults did not meet the aerobic and muscle strength guidelines as of 2016. These sedentary behaviors impact others through higher health insurance costs, lower worker productivity, and increased government spending on Medicare and Medicaid. As such, employers and policymakers have increasingly made use of financial incentives to improve wellness through exercise. Part of the success of such endeavors rests on the ability for incentives to change habits. While incentive programs have been shown to have strong positive short-run effects for problematic health behaviors, they have been less effective at generating long-lasting exercise habits, with the effects often disappearing once the incentive programs end (Royer et al. 2015). Therefore, understanding habit formation is of key importance for the design of incentive programs with longer-lasting effects.

In this paper, we analyze habit formation in exercise by examining the interdependence between past, present, and future exercise. We ask if exercise is habit-forming, in the sense that past exercise has a causal effect on current exercise. We also ask if individuals are forward-looking, in the sense that they make current exercise decisions depending on how much exercise they anticipate doing in the future. For example, an individual anticipating a busy work schedule tomorrow may choose to increase the amount of exercise today to make up for tomorrow's lost exercise. Our empirical strategy follows Becker, Grossman, and Murphy (1994) by considering the interaction between past and current exercise in a model with utility-maximizing individuals. The model is based on the assumption that past consumption of some goods influences their current consumption by affecting the marginal utility of current and future consumption. In relation to exercise, we hypothesize that higher levels of past exercise stimulate current exercise by increasing the marginal utility of current exercise more than the current marginal cost of the exercise.

Using login record data from a gym located in the Midwest, along with Quality Controlled Local Climatological Data from the National Oceanic and Atmospheric Administration (NOAA), we test our predictions by considering the response of gym attendance to snowstorms. Given that the Midwest is a region affected by moderate to severe winter snowstorms that often cause road closures, making it difficult for gym members to get to the gym, we examine whether past and future snowstorms lower gym attendance. The empirical results indicate that snowstorms have a highly significant effect on current gym attendance. In light of these results, we use snowstorm variation as an instrumental variable to estimate the dynamic relationship between past, present, and future exercise. Our results indicate that past and future exercise behaviors have an influence on present habits, implying rational habitual behavior in exercise and illustrating the importance of the intertemporal linkages in exercise behavior.

B. *Literature Review*

Our approach is motivated by a sizable and vibrant literature, beginning with Becker and Murphy (1988), aimed at studying rational addictive behavior. According to Becker and Murphy (1988), rational consumers maximize utility from stable preferences as they try to anticipate the future consequence of their actions. While the marginal utility of an addictive good in the current period depends on the consumption of the good in previous periods, rational consumers are farsighted, or forward-looking, in the sense that they anticipate the expected future consequences of their actions. In other words, current period consumption of an addictive good depends on past *and* future consumption levels of the good.

Although we tend to think of addictive goods in terms of substance abuse, people get addicted not only to alcohol, drugs, and cigarettes, but also to work, eating, music, television, their standard of living, other people, religion, and many other activities (Becker and Murphy 1988). While the degree of addictiveness varies from activity to activity and person to person, habits such as smoking, drinking, eating, and a host of others often meet two conditions required for addiction: reinforcement, in that the more you partake of the activity, the more you want to partake; and tolerance, in that the more that you partake of the activity, the lower your future utility given the amount of future consumption (Gruber and Koszegi, 2001).

While previous empirical studies analyzing tobacco consumption and voting behavior (Becker, Grossman, and Murphy 1994; Gruber and Koszegi 2001; Fujiwara, Meng, and Vogl 2015) have produced evidence supporting the theoretical framework of rational addiction proposed by Becker and Murphy (1988), fewer studies have focused on the implications of habit formation and rational addiction for exercise habits. Our research addresses this gap. Furthermore, our contribution to the literature on habit formation and rational behavior will inform business leaders and policy makers in dealing with rising healthcare costs by providing new insights for understanding whether incentive programs can be designed to change exercise habits and generate longer-lasting effects, which is an important open question in economics.

Following Becker and Murphy (1988), we explore the habitual nature of exercise by examining whether exercise exhibits adjacent complementarity. This happens if and only if past exercise influences current exercise by affecting the marginal utility of current and future exercise. In other words, we say there is adjacent complementarity in exercise if greater levels of past exercise stimulate current exercise by increasing the marginal utility of current exercise more than the present marginal cost. For example, engaging in exercise now may bring both monetary expenses (gym memberships) and high effort costs, but investing in improving physical condition today may improve the future enjoyment of the exercise (feeling healthier). Thus, adjacent complementarity in exercise implies that the additional gains in utility from engaging in more exercise today will outweigh the overall costs of future exercise (e.g., giving up watching an hour of television, or incurring gym memberships costs) over time. Thus, studying exercise habits has important implications for understanding time-inconsistent preferences stemming from both present consumption and projection bias.

Faced with high initial costs to change habits and long-run future benefits, an individual with present-biased preferences, or high preferences for goods involving immediate payoffs, may procrastinate on making healthy changes in behavior. Similarly, an individual with projection bias, meaning that the individual exaggerates the extent to which his/her future tastes will resemble his/her current tastes, may not appreciate that the costs of exercise (effort) are likely to fall over time and therefore may underinvest in establishing an initially difficult habit, such as exercising regularly (Royer et al. 2015). This renders the standard exponential discounting model, in which people discount the future at a constant fraction when comparing any two consecutive time periods, impractical for modeling exercise behavior of individuals with both present consumption and projection-biased preferences. A more realistic approach for modeling such behavior is presented in DellaVigna (2007), which highlights that individuals with self-control problems stemming from present consumption bias make consumption choices based on non-standard, time-inconsistent preferences modeled by

(1) $$U_t = u_t + \beta\delta u_{t+1} + \beta\delta^2 u_{t+2} + \beta\delta^3 u_{t+3} + \cdots$$

In (1), δ is the discount factor; u_t is the per-period utility of consumption; U_t gives the overall utility at time *t;* and $0 < \beta < 1$ shows discounting being steeper between the present and the immediate future, relative to discounting between any two periods in the distant future. This model captures the self-control problem driven by time-inconsistent preferences. The parameter β emphasizes the idea that, when evaluating outcomes between two time periods in the distant future, individuals are patient and have optimistic expectations about exercising. However, when comparing the present with the immediate future period, the discounting gets steeper, present bias becomes stronger, and individuals increase their consumption of leisure goods (e.g., watching TV) and delay painful tasks, such as exercise. Using this model, we can analyze how individuals make future consumption projections about investing in exercise. Individuals expect to consume an investment good in the future if

(2) $$\beta\delta b_1 + \beta\delta^2 b_2 \geq 0$$

Or

(3) $$b_1 + \delta b_2 \geq 0$$

In (2) & (3), $b_1 < 0$ represents utility derived from the good in the first period (since the good requires effort in this period), and $b_2 > 0$ represents the utility in the second period (since the good delivers a reward in this period), and β cancels out in (3), since comparison is taking place between two future periods and hence there is no present bias.

In the present period, however, the individual consumes an investment good if

(4) $$b_1 + \beta\delta b_2 \geq 0$$

In (4), b_1 is the utility derived from the good in the present period, and b_2 is the utility in the following period. In contrast to the overly confident projected outcomes of the individual given by (3), the parameter β in (4) does not cancel out, since comparison is taking place between the present and subsequent period. This denotes the steep discounting element in (1), while highlighting the essence of present consumption bias. Furthermore, the model emphasizes that when comparing the desired consumption with the actual consumption of an investment good, the individual consumes too little investment. The model also reiterates another important point discussed in Royer et al. (2015). Individuals with projection bias may fail to realize that the cost of exercise, b_1 in (4), is likely to fall when making comparisons between the present and subsequent periods in the future, while the reward of exercise, b_2 in (4), is likely to rise due to adjacent complementarity discussed in Becker and Murphy (1988). Thus, the individual with projection bias fails to appreciate the diminishing marginal cost (e.g., pain) of exercise and the respective increasing marginal utility (e.g., reward) as time moves forward.

2. Theoretical Discussion

Both projection bias and non-rational habitual behavior present a mechanism through which individuals may fail to make optimal intertemporal consumption decisions. Non-rational (myopic) individuals fail to have a consistent plan to maximize utility over time (Becker and Murphy 1988) and fail to anticipate the future consequences of their choices (current consumption is only affected by past consumption). In contrast, rational addicts are forward looking in that they make current consumption decisions considering both past and future consumption. A smoker, for example, would reduce consumption of cigarettes today in response to an anticipated future price increase in cigarettes. In relation to exercise, individuals anticipating higher exercise in a future period due to good weather may choose to increase their amount of current exercise. On the other hand, individuals anticipating less exercise in a future period due to a snowstorm may choose to decrease their amount of current exercise. We hypothesize that individuals are rational when it comes to exercise decisions and thus experience a future anticipation effect on current exercise. To test this hypothesis, we first explore a purely myopic empirical framework, in the case in which individuals ignore the future when making exercise decisions and thus are purely backward looking (myopic). We then examine a rational-choice model of exercise behavior and compare our results against the purely myopic empirical strategy in order to determine if exercise behavior is habit-forming in a purely backward looking (myopic) sense, or if habit formation in exercise is rational.

3. Data and Empirical Strategy

Gym Records and Climate Data

We merge data on daily gym member login records at a gym located in the Midwest region of the U.S. with corresponding daily Quality Controlled Local Climatological Data (QCLCD), provided by the National Centers for Environmental Information (NCEI), on precipitation and snowfall (in total daily mm) based on the gym's geographical location. The data runs from 1999 through 2009. NCEI is the world's largest provider of weather and climate data and provides the most comprehensive hourly, daily, and monthly summaries for approximately 1600 U.S. locations. The weather dataset includes a nearly complete set of observations on snowfall for the years included in our study, with only two missing observations that we drop from our sample. We restrict our gym record observations on individual gym visits to a one-visit maximum per individual per day, since individuals sometimes choose to split their daily exercise routine into two time sessions per day and accounting for these extra visits will likely bias our estimates. For example, we may be interpreting a change from a one-hour workout session during a single gym visit to two half-hour workout sessions at two different times in the day, which may result from an exercise schedule change due to changes in a person's daily work schedule, as *more* exercise. Since our study focuses on using snowstorm variation as an instrument for studying only the intertemporal linkages between past, current, and future exercise patterns, and we are able to control for individual, year, and week fixed effects in our models, the data is suitable for the purpose of our study. The Midwest is also an ideal location for this analysis, given its winter storm susceptibility, with snowstorm magnitudes ranging from moderate to severe.

Table 1 provides average daily snowfall by month in total millimeters. We use data at the weekly level to assess the effects of gym visits during a snowstorm week on gym visits in the subsequent week. Therefore, we drop observations from months with no snowstorms (May – September).

Table 1 - Average Daily Snowfall By Month In Total Millimeters

Month	**Mean**	**SD**	**Min**	**Max**
January	17	29	0	231
February	15	32	0	277
March	10	24	0	274
April	2	8	0	216
May	0	0	0	0
June	0	0	0	0
July	0	0	0	0
August	0	0	0	0
September	0	0	0	0
October	0	0	0	25
November	3	9	0	191
December	13	33	0	259

Figure 1 presents a time series of the probability of attending the gym each week around a 10-week snowstorm window, relative to the snowstorm week. **Figure 1** illustrates that gym attendance tends to fall one week before the snowstorm, and this declining effect persists even one week after the snowstorm.

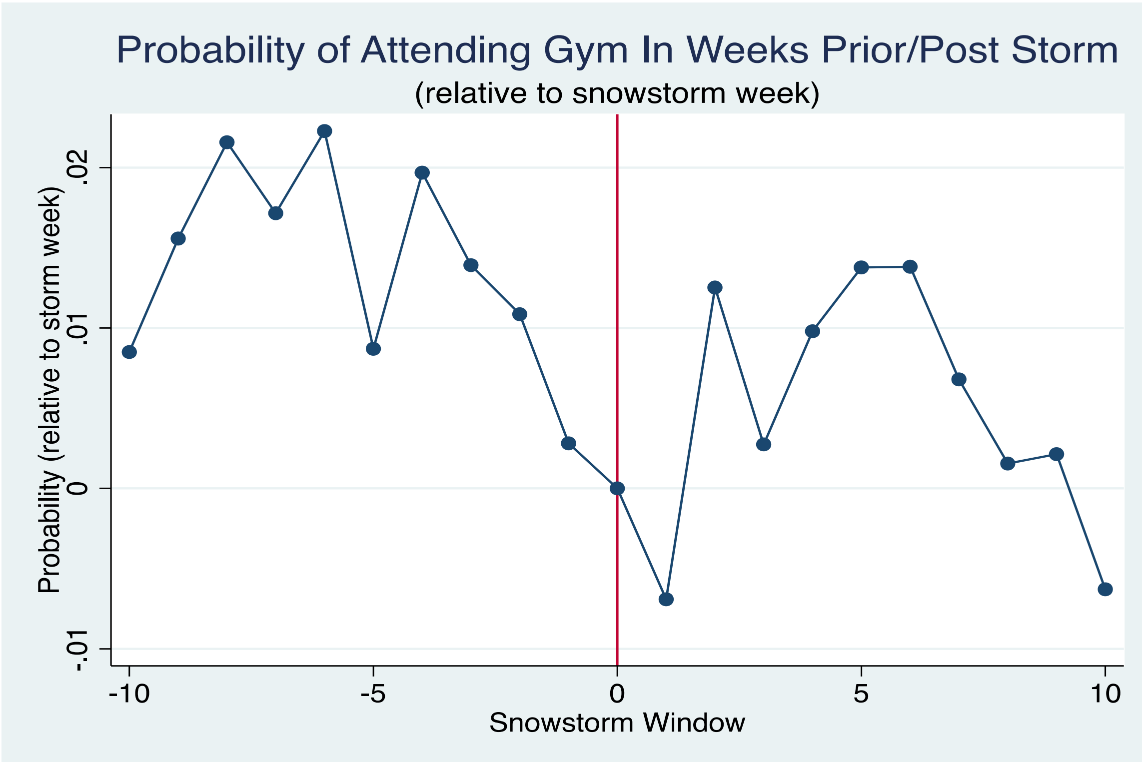

Figure 1 – Probability of Attending Gym In Week Prior/Post Snowstorm

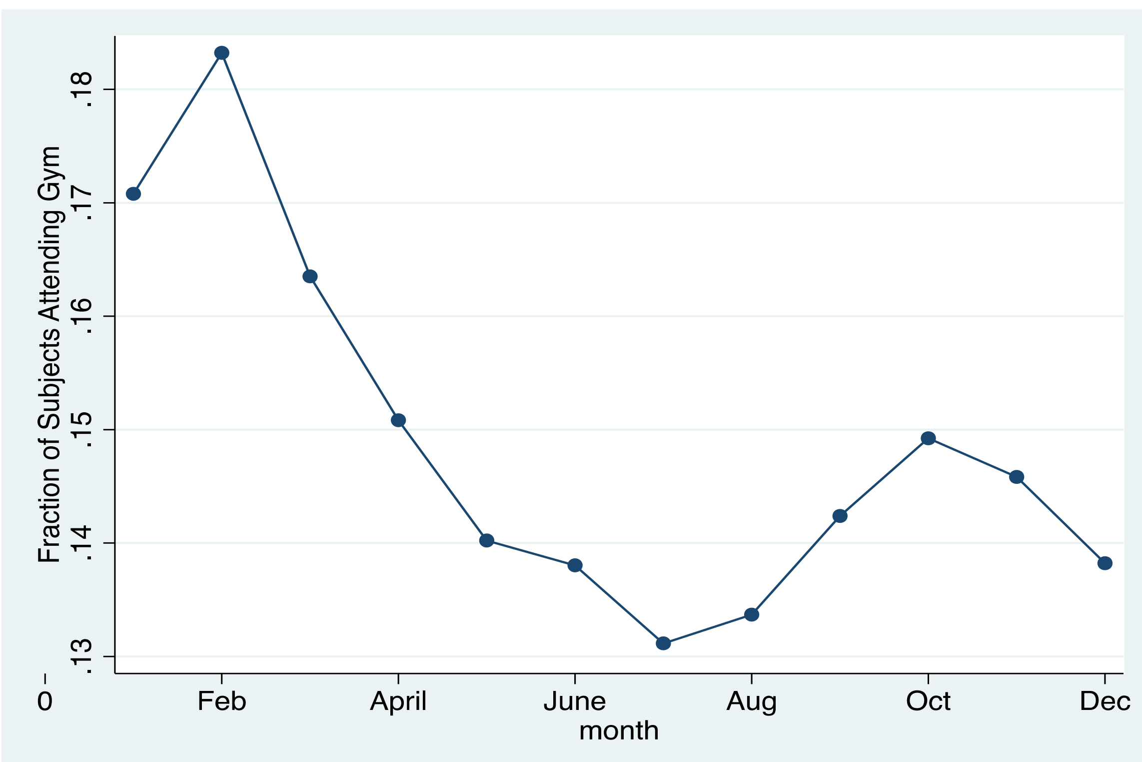

Figure 2 – Fraction of Total Gym Members Attending Gym By Month

Figure 2 graphs the fraction of total gym members attending the gym per month. **Figure 2** shows a monthly trend in gym attendance. We see that gym visits are highest during January and then follow a declining pattern ending with the month of July and beginning again in the month of October. We account for these trends in our regression models by using week of the year fixed effects.

4. Empirical Strategy and Results

Empirical Strategy

To address our research question of whether exercise behavior is myopic, we follow Becker, Grossman, and Murphy (1994) and present an empirical strategy in which we analyze exercise in week *t* as a function of exercise in the previous week *t* – 1. We present a myopic model of habitual behavior, where past exercise stimulates current exercise, but individuals ignore the future when making exercise decisions. The model takes the following form:

$$Visit_{it} = \theta_0 + \theta_1 Visit_{i,t-1} + \tau_y + \pi_w + \varpi_i + \varepsilon_{it} \quad (5)$$

In (5), the independent variable $Visit_{i,t-1}$ represents the total number of visits to the gym by an individual in week $t-1$, and τ_y, π_w, ϖ_i represent year, week and individuals fixed effects, respectively. Ordinary-least-squares estimation of equation (5) would lead to inconsistent estimates of the parameter of interest θ_1. The independent variable $Visit_{i,t-1}$ is likely to be endogenous, since unobservable individual characteristics and unobservable events affecting past gym visits may also affect current visits (omitted variables). Fortunately, a two-stage least squares estimator, using snowstorm variation in the years 1999-2009 as an instrument to identify θ_1, can address this concern. We use an indicator variable $Snowstorm$ in week $t-1$, taking on a value of 1 if there is a snowstorm in week $t-1$ and 0 otherwise, as the instrument. We define a snowstorm in terms of total daily snowfall, with 100 mm of daily snowfall indicating a snowstorm. We are confident in the validity of the instrument, since past snowstorms only affect current exercise through past gym visits. Furthermore, we base instrument relevance on the difficulty of commuting to and from the gym during a heavy snowstorm, given that roadblocks and other physical road constraints are quite common during severe snowstorms. The first stage regression relates gym visits in week $t-1$ to snowstorms in week $t-1$. The model takes the following form:

(6) $$Visit_{i,t-1} = \psi_0 + \psi_1 Snowstorm_{i,t-1} + \kappa_y + \nu_w + \mu_i + \rho_{it}$$

In (6), $\kappa_y, \nu_w,$ and $\mu_i$represent year, week, and individual fixed effects.

To address our research question of whether individuals are rational, or forward-looking, in their exercise decisions, we again use insights from Becker, Grossman, and Murphy (1994) and present a rational-choice model of exercise behavior, in which individuals try to anticipate the future consequences of their exercise choices and thus current exercise also depends on future exercise expectations. The model takes the following form:

(7) $$Visit_{it} = \beta_0 + \beta_1 Visit_{i,t-1} + \beta_2 Visit_{i,t+1} + \alpha_y + \varrho_w + \vartheta_i + \varpi_{it}$$

Similar to equations (5), we instrument the independent variables $Visit_{i,t-1}, Visit_{i,t+1}$ in (7) with snowstorm variation, this time using a one period lag and a one period lead of snowstorm as instruments. We use the Akaike information criteria (AIC) to decide on the amount of lags and leads to incorporate into our regressors in equations (5) and (7), which also corresponds to the amount of lags and leads that we include in our regressors from the first stage regressions outlined in equations (6), (8), and (9). The first stage regressions in the rational framework relate gym visits in week $t-1$ and week $t+1$ to snowstorms in week $t-1$ and week $t+1$. The models take the following form:

(8) $$Visit_{i,t-1} = \Psi_0 + \Psi_1 Snowstorm_{i,t-1} + \Psi_2 Snowstorm_{i,t+1} + \varkappa_y + \ell_w + j_i + \mathcal{E}_{i,t-1}$$

(9) $$Visit_{i,t+1} = \Upsilon_0 + \Upsilon_1 Snowstorm_{i,t-1} + \Upsilon_2 Snowstorm_{i,t+1} + \propto_y + \lambda_w + ℶ_i + \in_{i,t+1}$$

In (8) and (9), $\varkappa_y, \ell_w, j_i, \propto_y, \lambda_w, ℶ_i$, represent year, week, and individual fixed effects.

Results

Similarly to Becker, Grossman, and Murphy (1994), we begin our empirical estimation with the myopic model in equation (5). This model implies that the coefficient on instrumented future exercise β_2 in equation (7) should be zero, since individuals are purely backward looking in their current exercise decisions and future snowstorms and exercise changes have no impact on current exercise. We relax this assumption in the rational-choice model by examining whether anticipated future exercise is a significant predictor of current exercise, meaning that individuals are rational, or forward looking, in their present exercise decisions. We begin by performing two-stage-least-squares (TSLS) estimation using the myopic regression framework presented in equation (5), and then perform two-stage least-squares estimation using the rational exercise, or forward-looking, approach outlined in equation (7) to determine whether the respective parameter estimate for β_2 is statistically different from zero.

Table 2 presents our TSLS estimates for the first stage regression in the myopic exercise behavior model. The independent variables consist of past snowstorms ($Snowstom_{t-1}$) plus the other exogenous explanatory variables in the model. The dependent variable represents past gym visits ($Visit_{t-1}$). Columns (1) and (2) in **Table 2** present the results for all gym members, while columns (3) – (4) present results for regular exercisers (subjects attending gym at least once/week for the past month), and columns (5) – (6) present results for irregular exercisers (subjects not attending gym at least once/week for the past month). We analyze heterogeneity among regular and irregular exercisers because we expect these groups to behave differently. Thus, we are interested in seeing how sensitive these different groups are to exercise routine interruptions that can arise through shocks in a person's environment, such as snowstorms. As a robustness check for our coefficient estimates in **Table 2**, we include two regressions for each category of exerciser, one including only individual fixed effects and the other adding on week and year fixed effects. For the most part, we see that our coefficient estimates are sensitive to the inclusion of week and year fixed effects. Our coefficient estimates in **Table 2** provide a fair indication of instrument relevance, with most coefficient estimates being significant at the 1 percent level. These results are also intuitively pleasing, since they indicate that past snowstorms have a negative effect on past gym visits (for the models controlling for week, year, and individual fixed effects).

Table 2 – Two Stage Least Squares Regression Results (first stage)
Effect of Past Snowstorms on Past Gym Visits (myopic framework)

Dependent variable *Visit is total number of gym visits in a given week. Independent variable Snowstorm is 0/1 indicator=1 if there is a snowstorm in a given week and 0 otherwise*

	First Stage Regression Estimates					
	Overall		Regulars		Irregulars	
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$
$Snowstorm_{t-1}$	0.003	-0.037***	0.033***	-0.041***	-0.013***	-0.029***
	(0.005)	(0.004)	(0.009)	(0.009)	(0.004)	(0.004)
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Week Fixed Effects	No	Yes	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes	No	Yes
Observations	486,808	486,808	130,080	130,080	356,728	356,728
R-squared	0.380	0.389	0.326	0.335	0.195	0.205

Robust standard errors clustered by individual in parentheses. Columns (1)-(6) give two-stage least squares estimates from the first stage regressions in the myopic framework. *** p<0.01, ** p<0.05, * p<0.1

Table 3 presents our TSLS estimates for the second stage regression in the myopic framework, with $Visit_{t-1}$ treated as endogenous. The instruments used in **Table 3** consist of the past snowstorms ($Snowstom_{t-1}$) variable from the first stage plus the other exogenous explanatory variables in the model. Similar to Table 2, columns (1) and (2) in **Table 3** present the results for all gym members, while columns (3) – (4) present results for regular exercisers, and columns (5) – (6) present results for irregular exercisers. According to our parameter estimates of the myopic models presented in **Table 3**, the positive past exercise coefficients are consistent with our hypothesis that exercise is habit forming, in the sense that increases in past exercise have a positive effect on current exercise. However, none of the results are highly significant, with only the estimate in column 2 being significant at the 10 percent level.

Table 3 – Two Stage Least Squares Regression Results (second stage)
Effect of Past Gym Visits on Current Gym Visits (myopic framework)

Dependent variable Visit is total number of gym visits in a given week

	Second Stage Regression Estimates					
	Overall		Regulars		Irregulars	
Independent Variables	(1) $Visit_t$	(2) $Visit_t$	(3) $Visit_t$	(4) $Visit_t$	(5) $Visit_t$	(6) $Visit_t$
$Visit_{t-1}$	0.104	0.223*	0.102	0.104	0.229	0.253
	(0.148)	(0.127)	(0.214)	(0.196)	(0.187)	(0.166)
Individual Fixed Effects	No	Yes	No	Yes	No	Yes
Week Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	486,808	486,808	130,080	130,080	356,728	356,728
R-squared	0.020	0.392	0.019	0.323	0.020	0.289

Noes: Robust standard errors clustered by individual in parentheses. Columns (1)-(6) give two-stage-least-squares (2SLS) estimates with visit$_{t-1}$ treated as endogenous. The instruments in columns (1)-(6) consists of a one period lag of snowstorm plus the other explanatory variables in the model.
***** p<0.01, ** p<0.05, * p<0.1**

Table 4 presents our TSLS estimates for the first stage regression in the rational choice model for the effect of snowstorms on gym visits in the previous period, outlined in equation (8). The dependent variable ($Visit_{t-1}$) represents gym visits in the previous period. The independent variables of interest consist of past and future snowstorms ($Snowstorm_{t-1}$ and $Snowstorms_{t+1}$). We conduct a test of joint significance (F test) to test the hypothesis that the regressors corresponding to columns (1)-(6) have no joint explanatory power. In each case, our F statistic is well over 10, providing a fair indication of instrument relevance. These results are also intuitively pleasing, since they indicate that the presence of past snowstorm has a negative effect on past gym visits (for the models controlling for week, year, and individual fixed effects).

Table 4 – Two Stage Least Squares Regression Results (first stage)
Effect of Snowstorms on Past Gym Visits (rational framework)

Dependent variable *Visit is total number of gym visits in a given week. Independent variable Snowstorm is 0/1 indicator=1 if there is a snowstorm in a given week and 0 otherwise*

	First Stage Regression Estimates					
	Overall		Regulars		Irregulars	
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$	$Visit_{t-1}$
$Snowstorm_{t-1}$	0.005	-0.033***	0.034***	-0.040***	-0.012***	-0.027***
$Snowstorm_{t-1}$	0.039***	0.020***	0.027***	0.006	0.017***	0.005
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Week Fixed Effects	No	Yes	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes	No	Yes
Observations	486,808	486,808	130,080	130,080	356,728	356,728
R-squared	0.380	0.389	0.325	0.334	0.195	0.206

Columns (1)-(6) give two-stage least squares estimates from the first stage regressions in the rational framework. *** p<0.01, ** p<0.05, * p<0.1

Table 5 presents TSLS estimates for the first stage regression in the rational choice model for the effects of snowstorms on gym visits in the future period, only this time the dependent variable ($Visit_{t+1}$) represents gym visits in the future period. Similarly to **Table 4**, the independent variables of interest consist of past and future snowstorms ($Snowstorm_{t-1}$ and

$Snowstorms_{t+1}$). We conduct a test of joint significance (F test) to test the hypothesis that the regressors corresponding to columns (1)-(6) have no joint explanatory power. In each case, our F statistic is well over 10, once again providing a fair indication of instrument relevance.

Table 5 – Two Stage Least Squares Regression Results (first stage)
Effect of Snowstorms on Future Gym Visits (rational framework)

Dependent variable *Visit is total number of gym visits in a given week. Independent variable Snowstorm is 0/1 indicator=1 if there is a snowstorm in a given week and 0 otherwise*

	First Stage Regression Estimates					
	Overall		Regulars		Irregulars	
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables	$Visit_{t+1}$	$Visit_{t+1}$	$Visit_{t+1}$	$Visit_{t+1}$	$Visit_{t+1}$	$Visit_{t+1}$
$Snowstorm_{t-1}$	0.068***	0.022***	0.075***	0.038***	0.062***	0.024***
	(0.005)	(0.004)	(0.011)	(0.010)	(0.005)	(0.005)
$Snowstorm_{t+1}$	0.013***	-0.026	0.012***	-0.052***	-0.005***	-0.031***
	(0.005)	(0.005)	(0.012)	(0.011)	(0.004)	(0.005)
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Week Fixed Effects	No	Yes	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes	No	Yes
Observations	486,808	486,808	130,080	130,080	356,728	356,728
R-squared	0.369	0.389	0.310	0.323	0.287	0.297

Robust standard errors clustered by individual in parentheses. Columns (1)-(6) give two-stage least squares estimates from the first stage regressions in the rational framework.
***** p<0.01, ** p<0.05, * p<0.1**

Table 6 presents our TSLS estimates for the second stage regression in the rational framework, with the independent variables $Visit_{t-1}$ and $Visit_{t+1}$ (representing past and future gym visits) treated as endogenous. The instruments used in **Table 6** consist of the past and future snowstorms ($Snowstom_{t-1}$ and $Snowstom_{t+1}$) variables from the first stages (**Tables 4 and 5**) plus the other exogenous explanatory variables in the model. The dependent variable $Visit_t$ represents gym visits in the current period. Columns (1) and (2) present the results for all gym members, while columns (3) – (4) present results for regular exercisers, and columns (5) – (6) present results for irregular exercisers. While most of the coefficient estimates in **Table 6** are highly significant, they are also sensitive to the inclusion of week and year fixed effects. This may indicate that further inspection of model specification is needed.

Table 6 – Two Stage Least Squares Regression Results (second stage)
Effect of Past/Future Gym Visits on Current Visits (rational framework)

Dependent variable *Visit is total number of gym visits in a given week*

	First Stage Regression Estimates					
	Overall		Regulars		Irregulars	
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables	$Visit_t$	$Visit_t$	$Visit_t$	$Visit_t$	$Visit_t$	$Visit_t$
$Visit_{t-1}$	0.428***	1.162***	-0.094***	1.595***	0.706***	0.617***
	(0.022)	(0.399)	(0.483)	(0.409)	(0.296)	(0.293)
$Visit_{t+1}$	0.588***	1.178***	0.770***	1.091***	0.632***	0.403***
	(0.021)	(0.457)	(0.256)	(0.251)	(0.109)	(0.197)
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Week Fixed Effects	No	Yes	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes	No	Yes
Observations	486,808	486,808	130,080	130,080	356,728	356,728
R-squared	0.390	0.392	0.309	0.321	0.270	0.280

Robust standard errors clustered by individual in parentheses. Columns (1)-(6) give two-stage least squares estimates from the second stage regressions in the rational framework.
***** p<0.01, ** p<0.05, * p<0.1**

Conclusions

Previous research on habit formation has centered around studying the effects of past consumption of an addictive or habitual good on current consumption of the good, making assumptions of myopic consumption behavior. While rational choice models of forward-looking consumption behavior, following Becker and Murphy (1994), have also been present in the literature on habit formation, few have concentrated on the implications of rational addiction for exercise behavior. We address this gap by examining whether rational addiction models can be used to predict exercise habits. We ask if exercise is habit-forming, in the sense that past exercise has an effect on current exercise, but we take this analysis a step further and ask if individuals are forward-looking when making exercise decisions, meaning that their current exercise decisions are influence by how much exercise they anticipate to do in the future. We test whether individuals are purely myopic by identifying the effects of future exercise on current exercise using a rational-choice model of exercise behavior.

Our results indicate that exercise is habit-forming; with an increase in gym attendance by 1 visit in the current week translating to an increase of 1.16 visits to the gym in the following week. Furthermore, anticipating an increase in gym attendance by 1 visit in the following week leads to an increase in gym visits in the current week of 1.18 visits. This implies rational habitual behavior in exercise, with past and future changes in exercise significantly impacting current exercise. This evidence is inconsistent with the hypothesis that agents are myopic in regards to exercise decisions. Our contribution to the literature on habit formation and rational behavior in exercise will provide new insights for understanding how incentive programs can be designed to change exercise habits and generate longer-lasting effects. One caveat from our results is the inconsistency that we found in backward-looking behavior (past gym visits affecting current visits) between our purely myopic model and our rational exercise behavior model. Further research is needed to address this concern.

References

Becker, Gary S., Michael Grossman, and Kevin M. Murphy. "An Empirical Analysis of Cigarette Addiction. The American Economic Review, Vol. 84, No. 3. (Jun., 1994), pp. 396-418.

Becker, Gary S. and Kevin M. Murphy. (1988). "A Theory of Rational Addiction." *Journal of Political Economy 96(4): 675-700.*

Blumenkranz, Eric, Alan M. Garber, and Jeremy D Goldhaber-Fiebert. 2010. "Committing To Exercise: Contract Design For Virtuous Habit Formation. NBER Working paper No. 16624.

Carrera, Mariana, Heather Royer, Mark Stehr, and Justin Sydnor. (2014). "How Should We Structure Workplace Incentive Programs Targeting Exercise."

DellaVigna, Stefano. (2009). "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315-72.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl. (2015). "Habit Formation In Voting: Evidence from Rainy Elections." *American Economic Journal: Applied Economics.*

Gruber, Jonathan, and Botond Koszegi. (2001). "Is Addiction Rational? Theory and Evidence." *The Quarterly Journal of Economics.*pp. 2061-1303.

Royer, Heather, Mark Stehr, and Justin Sydnor. (2015). "Incentives, Commitments, and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company" *American Economic Journal: Applied Economics,* 7(3): 51-84.